Research Note

Check for updates

# Use of the Violence Risk Scale – Sexual Offense Version (VRS-SO) Within Sexually Violent Persons Evaluations

Sharon M. Kelley [1] (iD)

**[1]** *Sand Ridge Secure Treatment Center, Evaluation Unit, Madison, WI, USA.*

## Abstract

In many parts of the United States, individuals can be civilly committed as Sexually Violent Persons (SVP) to a secure treatment center based on their history of sexual offenses, current mental disorder, and current risk for sexual recidivism. While the specific criteria vary between jurisdictions, SVP civil commitment is indefinite, and periodic examinations occur to determine if ongoing commitment is necessary. Release recommendations may be made in part based on patients' treatment progress. Therefore, incorporating treatment change into periodic risk assessments is an important role of the SVP evaluator. The current paper sought to explore the benefits of using an actuarial tool within SVP populations to measure decreased sexual recidivism risk as a result of treatment change. Specific discussion of the use of the Violence Risk Scale – Sexual Offense version (Olver et al., 2007, https://doi.org/10.1037/1040-3590.19.3.318) is provided.

## Keywords

## Non-Technical Summary

### Background

In many parts of the United States, individuals can be civilly committed as Sexually Violent Persons (SVP) to a secure treatment center based on their history of sexual offenses, current mental disorder, and current risk for sexual recidivism. Individuals remain under commitment until they are determined to no longer have a mental disorder and/or have reduced their risk for sexual recidivism. Periodic court reports provide a review of the committed individual's progress and need for continued commitment. Release recommendations may be made in part based on the individual's treatment progress. Therefore, incorporating treatment change into periodic risk assessments is an important role of the SVP evaluator.

### Why was this study done?

Actuarial risk tools provide evaluators with a structured way of assessing risk factors based on research. These tools have been shown to be predictive in research studies. They have scoring rules, and the total score can be used to calculate estimated sexual recidivism rates based on the tool's norms. Some professionals have argued that actuarial measures of dynamic risk and treatment change cannot be used to assess future sexual reoffense risk for individuals who are civilly committed under SVP laws. The current paper challenges some of the common assumptions expressed by those who are of this view. The merits of one such measure, the Violence Risk Scale – Sexual Offense (VRS-SO) version are discussed in detail.

### What did the researcher do and find?

The present paper provides a review of methodological approaches and research related to assessing treatment change within SVP samples. Specifically, each of the methodological approaches are critically reviewed. Arguments for and against using the VRS-SO within SVP samples are also provided. I compared the high risk cases within three samples of the VRS-SO norms to cases at Sand Ridge (an SVP sample) who have VRS-SO scores. I looked at the average scores amongst the samples to see whether there were significant differences on initial risk scores and the treatment change score as provided by the VRS-SO. The results did find significant differences between some samples on risk assessment scores. The main difference was that the SVP sample was found to have higher VRS-SO scores on average (factors in need of treatment) at the beginning of treatment than the VRS-SO samples. However, the SVP sample demonstrated similar treatment change to most of the VRS-SO samples.

### What do these findings mean?

Overall, the research behind the VRS-SO is sufficient to meet admissibility requirements in most courts, and the results can provide important, meaningful information when evaluating individuals from high risk samples. Although the SVP sample had more treatment needs than those in prison settings, this is likely due to how SVP cases are selected. In other words, such cases are selected for potential SVP commitment because they have high levels of

treatment needs, frequently have not done well within prior treatment placements, and have demonstrated concerning behaviors. Importantly, these SVP cases were able to demonstrate an average amount of change. Being able to measure this change and determine whether the amount of treatment change an individual has made has sufficiently lowered their sexual reoffense risk is at the very core of SVP evaluations in determining when someone is ready for discharge.

**Highlights**
- The VRS-SO has been published in over 30 peer-reviewed journals.
- The VRS-SO is a frequently utilized actuarial measure of criminogenic needs by SVP evaluators, and second to the STABLE-2007.
- The VRS-SO has been validated to measure credible change, which has a risk-reducing effect.
- Within the normative samples, the VRS-SO includes 913 cases followed for 5- and 10-year periods. Of these, 82.3% of the cases were placed at high intensity treatment programs designed for those designated as high-risk.
- One civil commitment sample in the U.S. (i.e., Sexually Violent Person sample), demonstrated similar Static-99R and VRS-SO change scores as about half of the VRS-SO normative cases.
- Although the VRS-SO pre-treatment scores for the U.S. sample was one to two standard deviations larger than the VRS-SO norms, this may be due to how cases go through a selection process (e.g., screening, multiple evaluations, and formal trial) such that only the highest risk cases result in commitment.

In the United States (U.S.), individuals who demonstrate high levels of sexual violence risk can be civilly committed as "Sexually Violent Persons" (SVP) following the completion of their incarceration term. These laws are sometimes referred to as civil commitment or civil confinement laws. Currently, 20 states and the U.S. federal government have SVP laws. According to a recent survey in which 16 SVP sites responded, at least 4,948 individuals are currently committed under SVP laws in the U.S. (Schneider et al., 2019). While SVP laws vary between states, they generally require the respondent to be assessed as having (1) one or more convictions for qualifying sexual offenses; (2) a mental disorder that predisposes him/her to engage in future sexual offenses; and (3) a level of risk for future sexual offenses that exceeds the statutory threshold defined in the specific state. The last criterion is usually understood as the individual's risk being such that he/she is likely to sexually reoffend. Most states further interpret "likely" as "more likely than not" (Knighton et al., 2014). Interestingly, only six states define "more likely than not" as above a 50% threshold while the other states and federal system use clear and convincing evidence, have rejected a numerical threshold outright, or have not

**Psych**Open GOLD

yet defined what is meant by "more likely than not" (Knighton et al., 2014). Overall, SVP laws are aimed at the commitment of only the individuals assessed to be at the highest risk to sexually reoffend. Although SVP laws continue to engender much ethical debate, they have been upheld by the U.S. Supreme Court (*Kansas v. Crane*, 2002; *Kansas v. Hendricks*, 1997).

As SVP laws result in indeterminate commitment lengths, individuals' routes to discharge from a secure facility primarily requires them to demonstrate that their sexual recidivism risk has decreased. While aging within the secure facility is associated with a lowered risk for sexual re-offense (Helmus et al., 2012), this may be a long wait for those individuals who were committed in their 30s and 40s. SVP programs are tasked with providing treatment designed to reduce individuals' sexual recidivism risk, and treatment participation may be the most direct way for individuals to achieve this. Most SVP programs (over 90%) are designed to fit the Risk-Need-Responsivity model of treatment and provide Cognitive Behavioral Therapy treatment interventions (Andrews & Bonta, 2010; Schneider et al., 2019). It appears that the majority of civilly committed individuals are taking advantage of treatment opportunities (Schneider et al., 2019). Thus, assessing genuine treatment progress and other credible change indicators is an important role of the SVP evaluator as such positive changes may lead to recommendations for release.

Most SVP laws in the U.S. require the completion of court reports on a regular basis following an individual's initial commitment. The purpose of these reports is to provide the court with an update on the individual's treatment progress, whether they continue to meet the commitment criteria or whether they may be ready for supervised release and/or unconditional discharge. Considering whether the individual has made a sufficient amount of treatment change such that they have reduced their sexual recidivism risk is a critical aspect of these court reports. Yet, the methods for measuring treatment change varies greatly between programs and evaluators (Kelley et al., 2020; Schneider et al., 2019). A recent survey found that 36.1% of forensic evaluators (*n* = 90) believed that none of the available risk assessment measures had sufficient research support to measure treatment change, although the vast majority were using a structured measure to assess for treatment change (Kelley et al., 2020). Of those who were SVP evaluators (*n* = 59), 37.2% were not using any formal measure of dynamic risk, although when dynamic measures were selected, SVP evaluators tended to use the STABLE-2007 (Hanson et al., 2007) followed by the Violence Risk Scale – Sexual Offense version (VRS-SO; Olver et al., 2007). The current paper sought to explore the benefits of using an actuarial tool within SVP populations to measure decreased sexual recidivism risk as a result of treatment change.

# Methodological Options When Assessing Treatment Change

There appear to be two main avenues for evaluating risk and treatment change within SVP assessments: empirically-guided clinical judgment and the use of a structured tool to measure risk and treatment change. For the purposes of this paper, I define empirically-guided clinical judgment as identifying factors that have been found to be associated with sexual re-offense risk, but assigning one's own coding rules and level of importance to these risk factors (e.g., use of the Mann et al., 2010 study). Evaluators using this approach tend to assume risk is reduced once the individual completes or nears completion of a treatment program. This differs from a Structured Professional Judgment (SPJ) tool, which provides explicit coding rules in determining when a risk factor is present, although the evaluator remains free to place more importance on some risk factors than others when considering the overall risk. This also differs from an actuarial tool in which there are firm coding rules for each of the items, and the items are given numerical scores that are summed for a total score, which can be used to compare with existing norms (Kelley et al., 2020).

SVP evaluators using the first method use empirically-guided clinical judgment at multiple steps in the evaluation process: selecting which risk factors they want to include; choosing how to determine when the risk factors are present for a given case; determining what constitutes treatment change and when that treatment change is evident; determining the density of change such that the individual's risk can be expected to be lower within a specific area or lower overall; and determining how to integrate this treatment change with other assessments (e.g., Static-99R; Helmus et al., 2012) to obtain a final risk estimate. The problem is that methods employed at each of these steps, and even the behaviors considered to be evidence of change, can vary between evaluators or even between cases for the same evaluator. This is because each evaluator is using their own internal representation of what marks acceptable change behaviors. Evaluators also risk identifying factors that have little bearing on the prediction of sexual recidivism risk and may degrade predictive value (e.g., using risk factors from a research article that do not strongly correlate with sexual risk) (Cohen et al., 2020; Garb & Wood, 2019).

Evaluators using empirically-guided clinical judgment may assess change in one of two ways. Some may assume risk is reduced once the individual has been classified as having completed a treatment program (e.g., placement in the highest treatment phase or another designation symbolizing near or full treatment completion) without specifically quantifying the amount of risk reduction the individual achieved. Others may attempt to better capture this risk reduction by having a pre-assigned statistical weight for "treatment completion" (Bayesian approach; Elwood, 2018). There are only a few jurisdictions where individuals committed under SVP laws are discharged following the completion of treatment or determination they have mastered treatment concepts (Schneider et al., 2017). Therefore, most SVP evaluators are in a position to determine

PsychOpen GOLD

when a patient has received enough treatment through which they have adequately addressed their dynamic risk factors and lowered their risk. Use of an empirically-guided approach may not account for differences in density of dynamic risk factors at the outset of treatment or differences in meaningful change related to these risk factors.

Evaluators using a structured tool to measure treatment gains are either using an SPJ instrument or an actuarial measure of dynamic risk factors. Both types of tools help to focus the evaluator on risk-relevant information and anchor the evaluator's opinion. While SPJ tools have been successfully used in clinical settings, the emphasis on the evaluator's clinical judgment to determine the overall risk can result in the evaluator being overly influenced by a single risk factor or florid case detail (this can also occur when using empirically-guided clinical judgment). This is particularly problematic in adversarial SVP settings where research has demonstrated that evaluators' judgment is susceptible to potential bias (e.g., allegiance bias; Murrie et al., 2009; Murrie et al., 2013). Indeed, the evaluator assigned to the case may be more predictive of outcome than case factors themselves, and this is especially true when the assessment methodology requires more use of subjective judgment (Chevalier et al., 2015; Kahn, 2017; Murrie et al., 2009).

Actuarial tools better control use of clinical judgment and provide more sophisticated information (e.g., absolute recidivism rates), which helps minimize bias and increase predictive accuracy (Kelley & Thornton, 2015; Murrie et al., 2009; Olver et al., 2018). Importantly, actuarial tools are helpful in determining when an individual has demonstrated a sufficient level of treatment change that he has lowered his risk to sexually re-offend (i.e., has he actually benefitted from treatment?) (McGrath et al., 2012; Olver et al., 2018). The indeterminate nature of SVP laws may increase individuals' motivation to engage in impression management regardless of whether characterological traits may also be present. It can be reasonably expected that individuals will want to present well to obtain their liberties. This, however, presents some difficulties when it comes to assessing treatment gains. How can we discriminate between individuals who are only superficially engaged in treatment versus those who are meaningfully participating in treatment and internalizing the skills they are learning? Actuarial tools measuring dynamic risk factors are scored on the basis of documentation of observed behavioral change (Olver et al., 2018) and not simply advancement within a treatment program. Treatment change is examined based on the areas of need/dynamic risk factors at the start of treatment. In other words, the specific changes and the amount of change an individual needs to make depends on their initial dynamic risk (Olver et al., 2018). Without having a structured method to compare a client's current change with his pretreatment baseline, evaluators may not be able to accurately judge when the individual has received a sufficient amount of treatment.

PsychOpen GOLD

# Empirical Support for Actuarial Tools to Measure Treatment Change

There are multiple measures of criminogenic needs (i.e., dynamic risk) in circulation, including four actuarial measures: STABLE-2007, VRS-SO, Sex Offender Treatment Intervention and Progress Scale (SOTIPS; McGrath et al., 2012), and Sexual Risk Assessment – Forensic Version (SRA-FV; Thornton & Knight, 2015). The first three seek to measure sexual recidivism risk, which may change as a result of treatment interventions. The SRA-FV does not directly measure treatment change and evaluators using this scale have to consider such changes outside of the SRA-FV results. Although the factors within the SRA-FV are potentially changeable, the coding instructions for this tool focus on historical information as opposed to measuring treatment change.

For the purposes of this paper, I will focus on the use of the VRS-SO with an SVP population since I have direct experience with the tool and the majority of the SVP evaluators at the Sand Ridge Secure Treatment Center in Wisconsin are using the VRS-SO (66.7%). The VRS-SO was developed for use with adult, biologically-born males. It includes 17 dynamic items rated on a scale from 0 to 3. Scores of 2 and 3 are considered treatment needs and so are additionally rated as to the individual's readiness for change based on a modified transtheoretical stage of change model (Olver et al., 2007). Treatment change is measured based on the individual's movement across the stages of change since the beginning of treatment. Movement between each stage of change (with the exception of precontemplation to contemplation) results in a 0.5 change score. The change score can reflect treatment gains or the result of another credible change agent (larger change scores) as well as treatment regression (smaller change scores). Thus, individuals can move back and forth on the stages of change. The VRS-SO can be used by evaluators to generate an individual's pre-treatment dynamic risk score and treatment change score, both of which are combined with either the VRS-SO static score or Static-99R to obtain an integrated risk estimate (Olver et al., 2018). The integrated risk estimate reflects an individual's treatment change in relationship to their pre-treatment dynamic and static risk.

The VRS-SO is gaining increased popularity (Kelley et al., 2020). There are over 30 peer reviewed publications involving the VRS-SO (a reference list can be obtained from the author). The VRS-SO has undergone one independent validation study (Beggs & Grace, 2010, 2011), and has been studied in Canada, New Zealand, and Austria. In addition to the manual, the VRS-SO includes a User's Workbook available at https://psynergy.ca/vrs-so. The VRS-SO demonstrates good inter-rater reliability with Intraclass Correlation Coefficients (*ICC*) ranging from .73 to .97 for pre-treatment dynamic scores and .68 to .83 for change scores (Beggs & Grace, 2010, 2011; Holwell et al., 2017; Olver et al., 2007; Sowden & Olver, 2017). The VRS-SO has demonstrated good predictive validity within different samples including those composed of individuals identified as having pedophilic disorder and high levels of psychopathic traits (Eher et al., 2015;

PsychOpen GOLD

Olver et al., 2007; Sewall & Olver, 2019). Additionally, the VRS-SO has demonstrated good calibration for individuals identified as being "high risk." Specifically, Olver and Eher (2020) report an *E/O* index of 1.01 for individuals in the Well Above Average risk group. The pre-treatment dynamic and treatment change scores have both been found to have a unique incremental effect in the prediction of sexual recidivism risk beyond the Static-99R (Olver et al., 2018).

## Use of the VRS-SO to Measure Treatment Change Within SVP Cases

Evaluators who choose not to use actuarial tools frequently point to the SVP population as being substantially different than any of the research and normative samples generally supporting the use of an actuarial tool. Reasoned arguments include the higher risk of the SVP population, the increased intensity of dynamic risk factors, and the average length of confinement. The normative samples upon which actuarial risk measures are based include intensive treatment programs that typically do not exceed two years (Hanson et al., 2007; Hanson et al., 2015; Olver et al., 2020). When examining the SVP data in Wisconsin, the average length of stay from commitment to discharge was 10.5 years as of January 2020 ($N$ = 209, $SD$ = 5.3, $Mdn$ = 10.0 years). The high intensity treatment programs within VRS-SO normative samples generally lasted about eight to nine months and included 420 to 480 sex offense-specific treatment hours within that time (Olver et al., 2020). This can be translated into about 46.7 to 53.3 monthly sex offense-specific treatment hours within a nine-month period of time. Compare this with the average monthly treatment hours of an SVP program. According to the annual survey completed by the Sex Offender Civil Commitment Program Network (SOCCPN), the average monthly treatment dose for sex offense treatment groups is 16.9 hours with an additional 15.2 hours of psychoeducation groups (32.1 monthly treatment hours; Schneider et al., 2019). It would take 15 months to achieve 480 treatment hours or, if we exclude psychoeducation groups, over two years. Despite this, it is certainly the case that SVP sites have extremely long treatment stays.

It may also be true that the degree of pre-treatment static and dynamic risk is greater for SVP samples than those in the VRS-SO normative samples. This may also partially account for the longer time individuals remain in treatment. However, this would be expected within any high intensity treatment program, which is aimed to only target the highest risk individuals within the population. The actuarial risk of SVP samples may even exceed those of a high intensity treatment program within the prison system since individuals tend to be referred for SVP commitment due to continued concerns about risk, behavior management problems, prior community supervision failures, and unsuccessful treatment in a high intensity program while incarcerated. This "sifting" effect results in a specific type of sample with high static and dynamic risk, and which

is most likely composed of individuals that engage in treatment interfering behaviors. Yet, this should not preclude application of the VRS-SO in the SVP setting. Rather, such individuals will need to demonstrate larger amounts of change within treatment due to their higher levels of pretreatment risk, and they may take longer to make that treatment change due to their treatment resistant behaviors.

Given the low rate of discharges from SVP sites and lack of well-designed SVP recidivism studies, it has not been possible to validate any of the actuarial risk measures that specifically measure treatment change within an SVP population. In the absence of such studies, we can consider whether the tool's normative samples are different enough from an SVP sample that caution is warranted. The VRS-SO norms include 913 cases from four different samples. This includes one sample composed of low, moderate, and high intensity treatment programs (Olver et al., 2014), one sample that can be considered a moderate-high intensity program (Beggs & Grace, 2010), and two samples from a high intensity program (Olver et al., 2007; Sowden & Olver, 2017; see Olver et al., 2020 for more information). As noted in Figure 1, risk scores within the VRS-SO norms fall within a normal distribution with the majority of the cases receiving average risk scores but sufficient numbers of cases in the low and high risk ends such that recidivism estimates can be estimated from logistic regression modeling.

**Figure 1**

*Distribution of VRS-SO Total Scores Within the Normative Samples*
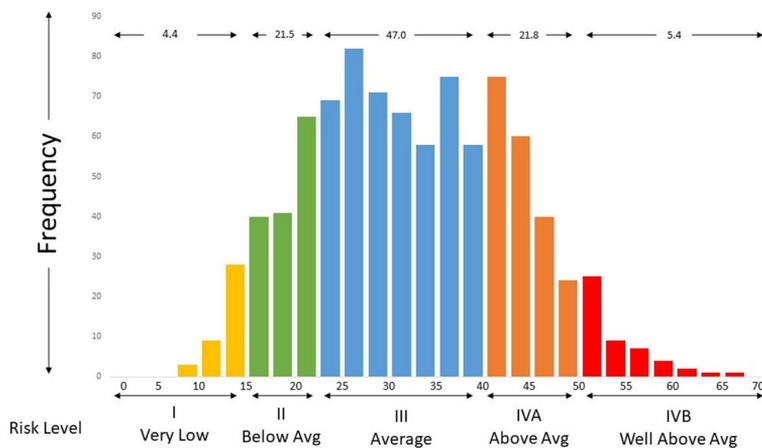


*Note.* Violence Risk Scale – Sexual Offense version (VRS-SO; Olver et al., 2007) total scores include the VRS-SO static and pre-treatment dynamic scores. From Olver, M. E. (2019, November). *Violence Risk Scale-Sexual Offense Version (VRS-SO) Training.* Workshop conducted at the 38th Annual ATSA Research and Treatment Conference, Atlanta, Georgia, United States. Reprinted with permission.

Olver (2018) reported that of the 913 cases in the norms, 755 cases were specifically from high intensity treatment programs. In order to directly analyze whether these cases are significantly different than typical SVP cases, I obtained the descriptive statistics for SVP cases with Static-99R and VRS-SO scores at Sand Ridge (Kelley, 2018) as well as the descriptive statistics from the three samples reported in Olver (2018).[1] A one-way analysis of variance (ANOVA) using sample sizes, means, and standard deviations was used to examine differences between the four high risk samples (Cohen, 2002). This approach was found to be preferable to comparing an average of the VRS-SO sample means to the SVP sample mean since there was found to be significant differences in variance between the three VRS-SO samples. The results are presented in Table 1.

Overall, the results indicate that the Clearwater 1 sample (Olver et al., 2007) had a significantly lower average Static-99R score than the National Sex Offender Program (NaSOP; Olver et al., 2014) and SVP sample. However, the average Static-99R score for the SVP sample was not significantly different than the Clearwater 2 (Sowden & Olver, 2017) or NaSOP samples. The NaSOP and Clearwater 2 samples did not have significantly different average pre-treatment dynamic and change scores. However, the means for both samples were significantly higher than the Clearwater 1 sample. The SVP sample was not significantly different than the NaSOP and Clearwater 2 samples with regards to average change scores, but it was different for average pre-treatment dynamic scores. This result is not unexpected and makes conceptual sense. As noted previously, the individuals who are committed to Sand Ridge may have failed to benefit from previous treatment interventions and/or are identified as likely candidates under the SVP law due to having high levels of static and dynamic risk. The VRS-SO helps to quantify how much change they need to make.

The cases within Clearwater 1 appear to have lower risk at the start of treatment. The average pretreatment dynamic score for the SVP sample was just over two standard deviations higher than Clearwater 1. However, the average pretreatment dynamic score for the SVP sample was just over one standard deviation higher than the NaSOP and Clearwater 2 samples, which make up just about half of the cases in the VRS-SO norms. As the SVP sample was within about a standard deviation of these samples, this indicates that there is sufficient overlap of cases with very high pre-treatment dynamic scores in the norms for logistic regression modeling of recidivism estimates in the VRS-SO normative samples to be relevant to SVP evaluations. There is no reason to indicate that the SVP scores in the current sample would not be similar to other SVP samples since evaluations examining commitment criteria usually seek to identify those with a high density of dynamic risk factors.

---

1) The data from (Kelley, 2018) and used in the current paper was based on summary descriptive data provided in internal program briefings and did not contain any identifying information. As such, review and approval was not required from the internal review board.

**Table 1**

*ANOVA Results of Risk Scores Between High Risk Samples*

| Risk tool | Samples (N = 937) | | | | F | $\eta^2$ | Tukey HSD Post-hoc test results | | | | | |
| | Clearwater 1 (n = 321) | NaSOP (n = 254) | Clearwater 2 (n = 180) | SVP (n = 182) | | | Clearwater 1 vs. NaSOP | Clearwater 1 vs. Clearwater 2 | Clearwater 1 vs. SVP | NaSOP vs. Clearwater 2 | NaSOP vs. SVP | Clearwater 2 vs. SVP |
| | M (SD) | | | | | | | | | | | |
| Static-99R | 4.6 (2.3) | 5.3 (2.1) | 4.9 (2.2) | 5.2[a] (1.7) n = 331 | 6.95*** | .02 | 0.7*** 95% CI [0.3, 1.2] | 0.3 95% CI [0.2, 0.8] | 0.6** 95% CI [0.2, 1.0] | -0.4 95% CI [-0.9, 0.1] | -0.1 95% CI [-0.5, 0.3] | 0.3 95% CI [-0.2, 0.8] |
| VRS-SO Pre-dynamic | 25.0 (7.5) | 31.8 (6.7) | 31.2 (5.4) | 39.4 (4.9) | 196.09*** | .39 | 6.8*** 95% CI [5.4, 8.2] | 6.2*** 95% CI [4.7, 7.8] | 14.4*** 95% CI [12.9, 15.9] | 0.6 95% CI [-2.2, 1.0] | 7.6*** 95% CI [6.0, 9.2] | 8.2*** 95% CI [6.5, 9.5] |
| VRS-SO change score | 2.6 (2.1) | 4.0 (2.9) | 4.2 (3.3) | 4.4 (3.0) | 23.67*** | .07 | 1.4*** 95% CI [0.8, 2.0] | 1.6*** 95% CI [0.9, 2.3] | 1.8*** 95% CI [1.1, 2.5] | 0.2 95% CI [0.5, 0.9] | 0.4 95% CI [0.3, 1.1] | 0.2 95% CI [0.6, 1.0] |

*Note.* Clearwater 1 refers to the Olver et al. (2007) sample. NaSOP or National Sex Offender Program refers to the Olver et al. (2014) sample. Clearwater 2 refers to the Sowden and Olver (2017) sample. SVP refers to a sample of patients at Sand Ridge Secure Treatment Center who received Violence Risk Scale – Sexual Offense version (VRS-SO) scores between 2014 and 2017.

[a]The average Static-99R total score for the SVP sample is the average for all patients committed at Sand Ridge in 2018 (*n* = 331).

** *p* < .01. *** *p* < .001.

# Conclusions

A primary role for SVP evaluators is to be able to discriminate between those who evidence credible change from those who do not. For those not using a formal assessment tool, evaluators are relying on empirically-guided clinical judgment, as it is unlikely SVP evaluators are still using unstructured clinical judgment anymore. Using empirically-guided clinical judgment to determine whether an individual has made a sufficient amount of treatment change is subject to error and bias (Walfish et al., 2012). While some evaluators may over-estimate treatment change by placing too much weight on behaviors in treatment groups that may mask underlying ambivalence about change (e.g., he gets along with staff and peers), other evaluators may under-estimate treatment change by requiring unnecessarily high standards or by incorrectly assuming that if the individual still has treatment tasks then he has not sufficiently lowered his risk. Within an SVP setting, even those patients who have sufficiently lowered their sexual risk continue to have treatment assignments until they are discharged by the court, so this would create circular reasoning. Further, patients only require the level of treatment intensity that meets their pre-treatment risk needs. This will not be the same for everyone. As such, assigning the same weight for all "treatment completers" will result in a notable loss of precision in one's risk estimates. As Olver et al. (2018) describes, the estimated sexual recidivism rate changes in response to the pre-treatment risk and amount of change an individual makes. Thus, it is possible that an individual may have exhibited less change than their peer but, because their pretreatment risk was initially lower, their estimated sexual recidivism risk remains lower.

Actuarial tools are increasingly being used to measure treatment change within SVP settings (Kelley et al., 2020). This popularity is likely related to both the availability of measures and training as well as a preponderance of research studies demonstrating their reliability and validity. The VRS-SO has also been shown to effectively estimate changes in sexual recidivism risk as a result of treatment change in multiple settings (Olver et al., 2018). I conclude that there is sufficient research to support the use of an actuarial tool to measure treatment change with SVP populations. Specifically, the VRS-SO has been shown to be reasonably accepted within the field (Kelley et al., 2020); it has undergone multiple publications in peer review journals; it has a manual and a known error rate (Olver et al., 2020); and it has been demonstrated to have reliability as well predictive and incremental validity (Olver et al., 2018; Olver et al., 2020). In addition, it has been validated for use with challenging populations usually seen within SVP samples such as those diagnosed with pedophilic disorder or who have been identified as having a high level of psychopathic traits (Eher et al., 2015; Olver et al., 2007; Sewall & Olver, 2019).

The risk scores within the VRS-SO norms fall within a normal distribution so that both low risk scores and very high risk scores are represented within the norms. The norms are comprised of two high risk samples and a third sample containing additional high risk cases. This results in 755 cases specifically from high intensity treatment

programs (82.3% of the norms). The Static-99R and VRS-SO change scores within the SVP sample were comparable to two of the three high risk samples in the VRS-SO norms, and within one standard deviation of the third sample. The VRS-SO pre-treatment dynamic scores for the SVP sample were about one standard deviation from two of the VRS-SO normative samples (about half of the normative cases). This indicates that the sexual recidivism estimates derived from logistic regression modelling can be applied to SVP cases since there are a sufficient number of cases within the VRS-SO norms that have elevated risk scores, and the SVP sample was not found to have risk scores that are extreme deviations from the VRS-SO norms (e.g., a difference of three standard deviations). Logistic regression modeling does not require the SVP sample to have identical risk means as the normative samples, just that a sufficient number of higher risk cases are included. As with any measure, as the scores for an individual SVP case become unusually extreme, the more likely for error, which is accounted for by the widening confidence intervals provided within the results of the risk estimate.

The primary difference between the SVP and VRS-SO samples being related to pre-treatment dynamic risk should be an unsurprising finding since SVP commitment tends to occur after previous interventions have failed (e.g., poor compliance with community supervision and/or prior treatment interventions). While individuals committed as SVPs have extensive treatment needs and spend many years institutionalized, this does not preclude the use of the VRS-SO. Rather, it is all the more critical that treatment change be captured in a valid and reliable manner. Asserting that the VRS-SO cannot be used with an SVP population because SVP patients have higher risk than the people in the normative samples is an inaccurate argument given that there are a sufficient number of high-risk cases with pre-treatment dynamic scores that overlap those in an SVP sample. While SVP patients have a higher level of risk and more time in treatment, this allows them ample opportunity to make large treatment gains. Further arguing that the VRS-SO cannot be used due to a lack of validation research within the SVP population is an unnecessarily conservative argument, which would also preclude the use of most risk tools, other evaluation methodologies, and treatment outcome studies. The VRS-SO has demonstrated reliability and predictive validity within heterogeneous samples. There is no substantive evidence to indicate that these results cannot be generalized to an SVP setting where the precise measurement of treatment change is critical for evaluation practice standards, community protection, and individual rights of the patients. Although the current results are not dispositive, they provide some initial support for using the VRS-SO within SVP evaluations. An important next step would involve an investigation to determine the predictive accuracy of the VRS-SO for those who have been discharged from SVP settings.

# References

Andrews, D. A., & Bonta, J. L. (2010). *The psychology of criminal conduct* (5th ed.). Cincinnati, OH, USA: Anderson.

Beggs, S. M., & Grace, R. C. (2010). Assessment of dynamic risk factors: An independent validation study of the Violence Risk Scale: Sexual Offender Version. *Sexual Abuse, 22*(2), 234-251. https://doi.org/10.1177/1079063210369014

Beggs, S. M., & Grace, R. C. (2011). Treatment gains for sexual offenders against children predicts reduced recidivism: A comparative validity study. *Journal of Consulting and Clinical Psychology, 79*(2), 182-192. https://doi.org/10.1037/a0022900

Chevalier, C. S., Boccaccini, M. T., Murrie, D. C., & Varela, J. G. (2015). Static-99R reporting practices in sexually violent predator cases: Does norm selection reflect adversarial allegiance? *Law and Human Behavior, 39*(3), 209-218. https://doi.org/10.1037/lhb0000114

Cohen, B. H. (2002). Calculating a factorial ANOVA from means and standard deviations. *Understanding Statistics, 1*(3), 191-203. https://doi.org/10.1207/S15328031US0103_04

Cohen, T. H., Lowenkamp, C. T., Bechtel, K., & Flores, A. W. (2020). Risk assessment overrides: Shuffling the risk deck without any improvements in prediction. *Criminal Justice and Behavior, 47*(2), 1609-1629. https://doi.org/10.1177/0093854820953449

Eher, R., Olver, M. E., Heurix, I., Schilling, F., & Rettenberger, M. (2015). Predicting re-offense in pedophilic child molesters by clinical diagnoses and risk assessment. *Law and Human Behavior, 39*(6), 571-580. https://doi.org/10.1037/lhb0000144

Elwood, R. W. (2018). Updating probability in sex offender risk assessment. *International Journal of Offender Therapy and Comparative Criminology, 62*(7), 2063-2080. https://doi.org/10.1177/0306624X17711880

Garb, H. N., & Wood, J. M. (2019). Methodological advances in statistical prediction. *Psychological Assessment, 31*(12), 1456-1466. https://doi.org/10.1037/pas0000673

Hanson, R. K., Harris, A. J. R., Scott, T., & Helmus, L. (2007). *Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision Project* (Corrections Research User Report No. 2007-05). Ottawa, ON, Canada: Public Safety Canada.

PsychOpen GOLD

Hanson, R. K., Helmus, L. M., & Harris, A. J. R. (2015). Assessing the risk and needs of supervised sexual offenders: A prospective study using STABLE-2007, Static-99R, and Static-2002R. *Criminal Justice and Behavior, 42*(12), 1205-1224. https://doi.org/10.1177/0093854815602094

Helmus, L., Thornton, D., Hanson, R. K., & Babchishin, K. M. (2012). Improving the predictive accuracy of Static-99 and Static-2002 with older sex offenders: Revised age weights. *Sexual Abuse, 24*(1), 64-101. https://doi.org/10.1177/1079063211409951

Holwell, M. J., Beggs Christofferson, S. M., & Olver, M. E. (2017). Evaluating the inter-rater reliability of the Violence Risk Scale – Sexual Offense version (VRS-SO) in a community-based treatment setting. *Sexual Offender Treatment, 12*(1), 1-14. Retrieved from http://www.sexual-offender-treatment.org/160.html

Kahn, R. E. (2017). Influence of case details and evaluator differences in SVP cases. In S. M. Kelley (Chair), *Evaluator struggles: Risk aversion, impotence, and decision making.* Symposium conducted at the 36th Annual ATSA Research and Treatment Conference, Kansas City, Missouri, USA.

Kansas v. Crane, 534 US 407 (2002).

Kansas v. Hendricks, 521 US 346 (1997).

Kelley, S. M. (2018, October). Protective factors and treatment change in high states evaluations. In A. J. R. Harris (Chair), *"Dynamic Smackdown." Dynamic Assessment – Promise or Pretext?* Symposium conducted at the 37th Annual ATSA Research and Treatment Conference, Vancouver, British Columbia, Canada.

Kelley, S. M., Ambroziak, G., Thornton, D., & Barahal, R. M. (2020). How do professionals assess sexual recidivism risk? An updated survey of practices. *Sexual Abuse, 32*(1), 3-29. https://doi.org/10.1177/1079063218800474

Kelley, S. M., & Thornton, D. (2015). Assessing risk of sex offenders with major mental illness: Integrating research into best practices. *Journal of Aggression, Conflict and Peace Research, 7*(4), 258-274. https://doi.org/10.1108/JACPR-02-2015-0162

Knighton, J. C., Murrie, D. C., Boccaccini, M. T., & Turner, D. B. (2014). How likely is "likely to reoffend" in sex offender civil commitment trials? *Law and Human Behavior, 38*(3), 293-304. https://doi.org/10.1037/lhb0000079

Mann, R. E., Hanson, R. K., & Thornton, D. (2010). Assessing risk for sexual recidivism: Some proposals on the nature of psychologically meaningful risk factors. *Sexual Abuse, 22*(2), 191-217. https://doi.org/10.1177/1079063210366039

McGrath, R. J., Lasher, M. P., & Cumming, G. F. (2012). The Sexual Offender Treatment Intervention and Progress Scale (SOTIPS): Psychometric properties and incremental predictive validity with Static-99R. *Sexual Abuse, 24*(5), 431-458. https://doi.org/10.1177/1079063211432475

Murrie, D. C., Boccaccini, M. T., Guarnera, L. A., & Rufino, K. A. (2013). Are forensic experts biased by the side that retained them? *Psychological Science, 24*(10), 1889-1897. https://doi.org/10.1177/0956797613481812

Murrie, D. C., Boccaccini, M. T., Turner, D. B., Meeks, M., Woods, C., & Tussey, C. (2009). Rater (dis)agreement on risk assessment measures in sexually violent predator proceedings: Evidence

PsychOpen GOLD

of adversarial allegiance in forensic evaluation? *Psychology, Public Policy, and Law, 15*(1), 19-53. https://doi.org/10.1037/a0014897

Olver, M. E. (2018, October). Dynamic risk assessment using the VRS-SO with indeterminate detention offenders. In A. J. R. Harris (Chair), *"Dynamic Smackdown." Dynamic Assessment – Promise or Pretext?* Symposium conducted at the 37th Annual ATSA Research and Treatment Conference, Vancouver, British Columbia, Canada.

Olver, M. E. (2019, November). *Violence Risk Scale-Sexual Offense Version (VRS-SO) Training*. Workshop conducted at the 38th Annual ATSA Research and Treatment Conference, Atlanta, Georgia, USA.

Olver, M. E., & Eher, R. (2020). Predictive properties and factor structure of the VRS-SO in an Austrian sample. *European Journal of Psychological Assessment, 36*(5), 748-757. https://doi.org/10.1027/1015-5759/a000551

Olver, M., Kelley, S., Johnson, L., & Wong, S. (2020). *Violence Risk Scale – Sexual Offense Version (VRS-SO): User's workbook*. Retrieved from https://psynergy.ca/vrs-so

Olver, M. E., Kingston, D. A., Nicholaichuk, T. P., & Wong, S. C. P. (2014). A psychometric examination of treatment change in a multisite sample of treated Canadian federal sexual offenders. *Law and Human Behavior, 38*(6), 544-559. https://doi.org/10.1037/lhb0000086

Olver, M. E., Mundt, J. C., Thornton, D., Beggs Christofferson, S. M., Kingston, D. A., Sowden, J. N., Nicholaichuk, T. P., Gordon, A., & Wong, S. C. P. (2018). Using the Violence Risk Scale – Sexual Offense version in sexual violence risk assessments: Updated risk categories and recidivism estimates from a multisite sample of treated sexual offenders. *Psychological Assessment, 30*(7), 941-955. https://doi.org/10.1037/pas0000538

Olver, M. E., Wong, S. C. P., Nicholaichuk, T. P., & Gordon, A. (2007). The validity and reliability of the Violence Risk Scale – Sexual Offender version: Assessing sex offender risk and evaluating therapeutic change. *Psychological Assessment, 19*(3), 318-329. https://doi.org/10.1037/1040-3590.19.3.318

Schneider, J. E., Jackson, R., Ambroziak, G., D'Orazio, D., Freeman, N., & Hebert, J. (2017, October). *SOCCPN Annual Survey of Sex Offender Civil Commitment Programs 2017.* Symposium conducted at the Annual SOCCPN Conference, Kansas City, Missouri, USA.

Schneider, J. E., Jackson, R., Ambroziak, G., D'Orazio, D., Freeman, N., & Hebert, J. (2019, November). *SOCCPN Annual Survey of Sex Offender Civil Commitment Programs 2019.* Symposium conducted at the Annual SOCCPN Conference, Atlanta, Georgia, USA.

Sewall, L. A., & Olver, M. E. (2019). Psychopathy and treatment outcome: Results from a sexual violence reduction program. *Personality Disorders, 10*(1), 59-69. https://doi.org/10.1037/per0000297

Sowden, J. N., & Olver, M. E. (2017). Use of the Violence Risk Scale – Sexual Offender version and the Stable-2007 to assess dynamic sexual violence risk in a sample of treated sexual offenders. *Psychological Assessment, 29*(3), 293-303. https://doi.org/10.1037/pas0000345

Thornton, D., & Knight, R. A. (2015). Construction and validation of SRA-FV Need Assessment. *Sexual Abuse, 27*(4), 360-375. https://doi.org/10.1177/1079063213511120

PsychOpen GOLD

Walfish, S., McAlister, B., O'Donnell, P., & Lambert, M. J. (2012). An investigation of self-assessment bias in mental health providers. *Psychological Reports, 110*(2), 639-644. https://doi.org/10.2466/02.07.17.PR0.110.2.639-644